Neurocomputing 207 (2016) 120–130

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

A sparse unmixing model based on NMF and its application in Raman image

Yifei Xu [a,d], Shuiguang Deng [a,c,*], Xiaoli Li [b,c], Yong He [b,c]

[a] College of Computer Science and Technology, Zhejiang University, PR China
[b] College of Biosystems Engineering and Food Science, Zhejiang University, PR China
[c] Cyrus Tang Center for Sensor Materials and Applications, Zhejiang University, PR China
[d] Department of Biosystems & Agricultural Engineering, Michigan State University, USA

ARTICLE INFO

Article history:
Received 10 November 2015
Received in revised form
25 February 2016
Accepted 18 March 2016
Communicated by Deng Cai
Available online 27 May 2016

Keywords:
Raman imaging
Sparsity
Plant cell wall

ABSTRACT

Spectral unmixing is a critical issue in multi-spectral data processing, which has the ability to determine the composition and the structural characteristics of the Raman image. Most of current unmixing methods work well to explore the materials in an ideal scenario. However, both the noise and the requirement of the prior knowledge limit their practical application. Thus, we propose a sparse method called to unmix spectra and apply it to explore the elucidate structural and spatial distribution of the plant cell well. GRSRNMF utilizes the blind source separation technology based on NMF to determine the basis elements (abundances) and their corresponding components (endmembers) without prior knowledge. GRSRNMF incorporates graph relationship and $L_{1/2}$ regularizer to improve the robustness and effectiveness. Besides, two proper indicators are designed to assess the unmixing method for Raman image when the standard spectrum library does not exist. Experiments are conducted on simulated datasets and the real-world Raman image to evaluate the performance of the proposed methods from various aspects. Experimental results illustrate that the proposed method favors sparsity and offers improved estimation accuracy compared to other methods.

1. Introduction

Raman imaging, as a new hyperspectral imaging[1–3], has been used to illustrate changes of molecular composition in a cell wall [4,5]. In the field of plant science and bioinformatics science, gaining each composition in a cell wall is a foundation work for further analysis. Unfortunalty, since the complex structure and interference between the components, it is still a critical issue to distinguish the in an efficient way. Traditional Raman image analysis methods make use of the peak intensity or peak integration to determine the components of the cell wall. However, these methods are helpless in case that some of the components are mixed. Recently, inspired by the hyperspectral unmixing, some efforts have been made to unmix Raman image applied in various fields [6–8], but there are still some challenges to be addressed: 1) pure components (endmembers) and their basis elements (abundances) are unknown. Moreover, for Raman image analysis, there is no standard spectrum library as a reference. Thus, it is required to unmix Raman image without any prior information and evaluate the performance with new indicators. 2) Similarity of the different materials with overlapped spectra limits the performance of the Raman unmixing. Hence, it is essential to improve the capacity of separating different materials.

To address the above challenges, we first propose a new unsupervised and sparse method called Graph-Regularized Sparse Recursive NMF (GRSRNMF) and two new metrics for spectral unmixing. Experiments are conducted on synthetic datasets and the real Raman image to evaluate its performance, and the results show that it outperforms other related methods on several metrics. As a summary, the main contributions can be listed as follows: 1) a sparse and unsupervised method based on NMF called GRSRNMF is proposed to unmix the components in a Raman image. Graph relationship and $L_{1/2}$ regularizer are incorporated into GRSRNMF to solve the overlapped spectra problem with a significant performance improvement. Compared with other state-of-art approaches, the proposed method can find the each component ideally without prior information, and cope with the situation that all the components are mixed in a random way. Through the complexity and sparsity analysis in Sections 3 and 4, we can find that our method works effectively in a reasonable

* Corresponding author at: College of Computer Science and Technology, Zhejiang University, PR China. Tel./fax: +86 571 88273422.
E-mail address: dengsg@zju.edu.cn (S. Deng).

http://dx.doi.org/10.1016/j.neucom.2016.03.063
0925-2312/

running time. 2) GRSRNMF determines the basis elements and their corresponding components in sequence, which is in the opposite order of the traditional hyperspectral unmixing method [9]. By doing so, the basis elements can verify the accuracy of the estimated components and vice versa. 3) Two proper metrics are designed when the standard spectrum library is absent to estimate the performance specified for Raman image.

The remaining of this paper is organized as follows: Section 2 gives an overview of the related work. Section 3 presents the detail of GRSRNMF. The datasets, parameter settings, and evaluation metrics are shown in Section 4. Section 5 presents the experiments and discusses the performance of GRSRNMF and the comparative methods. Section 6 concludes this paper and presents the future work.

Notations. The set of m-by-n real matrices is denoted $R^{m\times n}$; for $A\in \boldsymbol{R}^{m\times n}$, $\boldsymbol{A}_{i:}$ is the ith row of $\boldsymbol{A}$, $\boldsymbol{A}_{i:}$ is the ith column of $\boldsymbol{A}$ and $\boldsymbol{A}_{ij}$ is the entry at position (i,j); for $b\in \boldsymbol{R}^{m}$, we denote b_i as the ith entry of b. The set $R^{m\times n}$ with component-wise nonnegative entries is denoted $R_{+}^{m\times n}$; The L_0 of vector x denoted $\|x\|_0$ is the cardinality of the set $\{i|x_i\neq 0\}$; The L_2 of vector x denotes $\|x\|_2=\sqrt{\sum x_i^2}$; the L_1 of vector x is $\|x\|_1=\sum_i|x_i|$. The $L_{1/2}$ of vector x is $\|x\|_2=\sum_i|x^{1/2}|$; the Frobenius norm of matrix $\boldsymbol{A}$ is $\|A\|_F^2=\sum_{ij}a_{ij}^2$.

2. Related work

In general, unmixing methods work well under the assumption that the points conform to the Linear Mixture Model (LMM). Most of the current methods employ LMM to approximately describe the hypercube under Abundance Non-negativity Constraint (ANC) and Abundance Sum-to-one Constraint (ASC)[10]. Consequently, the task of unmixing problem is to factorize a high-dimension matrix into two low-dimensional matrices subjected to ANC and ASC. The algorithms for unmixing hypercube include Pixel Purity Index (PPI) [11], N-FINDR [12], Vertex Component Analysis (VCA) [13], Minimum Volume Enclosing Simplex (MVES) [14] and Automatic Target Generation Process (ATGP) [15]. These approaches work better only when assuming the pure pixels exists or at least p-1 (p stands for the number of endmember) spectral vectors in a pixel.

Since the assumption of pure pixels is not reliable, the blind source separation (BBS) approaches are developed from the perspective of statistics to deal with this issue [16]. Independent Component Analysis (ICA) explores an available transformation under the assumption that all the endmembers are statistical independent [17,18]. If the endmembers are independent, ICA provides the correct unmixing since the minimum of the mutual information corresponding to and only to independent sources [17]. Except for this demerit, the ANC constraint is also disabled in ICA. Besides, the methods based on Bayesian framework incorporate ANC and ASC directly in the prior distribution and parameter, and can provide a heuristic and global solution of hyperspectral unmixing problems as much as possible [19]. Typically, the hyperparameters and parameters in the joint posterior distribution resort to the sample generated algorithms-Markov chain Monte Carlo algorithm. The advantage of this sample method is that it can converge on the joint distribution over the parameters and hyperparameters. However, the computational complexity is still high despite that several available measures have been made to improve this situation [20,21].

As another significant BBS approach, Nonnegative Matrix Factorization (NMF) provides a part-based representation of the data, making the decomposition matrices more intuitive and interpretable [22–24], which is in accordance with the principle of hyperspectral image unmixing. However, NMF is disabled to separate all the components correctly due to the non-uniqueness of the solution [25,26]. For improving the performance of NMF, a variety of constraints is taken into account [27,28]. Nonnegative sparse coding incorporates sparse coding as sparsity constraint into NMF, whereas smoothness constraint is considered in [29]. These algorithms, however, are not designed for hyperspectral data analysis and, thus, fail to make full use of the characteristics of hyperspectral data, which compromises the performance when being applied into this domain. From a convex geometric point of view, minimum-volume-constrained NMF (MVCNMF) [30] utilizes the minimum volume constraint, which drives the virtual endmembers to enclose the data cloud. Except the above constraints, other constraints are taken into account to make the problem more well-posed, e.g., sparsity of the abundance matrix and piecewise smoothness of spectral signals [31], orthogonally [32], and sum-to-one the abundance [22]. Still, when the rank is modified, the methods based traditional NMF need to re-compute the solutions and become more time-consuming.

In this paper, we propose GRSRNMF, a new variant of NMF that overcomes the above demerits. As used in NMF, L_p regularizer is introduced to GRSRNMF to improve the performance. The L_0 regularizer means the number of zero elements in the basis elements matrix to yield the sparsest result. However, the solution of the L_0 regularizer is an NP-hard problem. The L_2 regularizer generates smooth but not sparse result. For L_1 regularizer, the sparsity property and its influence on the image unmixing have not been thoroughly investigated.

Therefore, we incorporate $L_{1/2}$ regularizer into GRSRNMF to enforce the sparsity of the abundance [33]. It is implemented through the multiplicative updating algorithm, which is an iterative application of a rescaled gradient descent. In virtue of sparsity, GRSRNMF leads to a more satisfactory result. In fact, all of the sparse constraints only consider the Euclidean structure of the data space that cannot be uniformed filled up by the hypercube. These hypercube data can be regarded as sampled data from or near a submanifold of an ambient space. Therefore, it is necessary to consider the intrinsic manifold structure while performing Raman image unmixing. Inspired by the manifold learning and sparse constraints, we then incorporate the manifold structure (graph relationship), leading to improved performance on several indicators.

3. Graph-Regularized Sparse Recursive NMF

3.1. GRSRNMF

In general, NMF can be described as follows: given a nonnegative input matrix $M\in \boldsymbol{R}_{+}^{m\times n}$ and an integer$r(1\leq r< \min(m,n))$, find two nonnegative metrics $U\in \boldsymbol{R}_{+}^{m\times r}$ and $V\in \boldsymbol{R}_{+}^{r\times n}$ be the solution of the following minimization problem.

$$\min_{\boldsymbol{V}\in \mathbf{R}^{r\times n},\, \boldsymbol{U}\in \mathbf{R}^{m\times r}} \|\boldsymbol{M}-\boldsymbol{U}\boldsymbol{V}\|_F^2 \quad \text{such that } \boldsymbol{U}\geq 0, \boldsymbol{V}\geq 0 \tag{1}$$

Due to the NP-hardness, practical algorithms cannot be expected to find provably optimal global solution in a reasonable time. However, with the help of the theorems Perron-Frobenius and Eckart-Young, it is easy to find a globally optimal rank-1 NMF in polynomial time. Specially, the Perron-Frobenius theorem implies that the dominant left and right singular vectors of a nonnegative matrix are nonnegative, while the Eckart-Young theorem states that the outer product of these dominant singular vectors is the best possible rank-1 approximation in the Frobenius norm.